DOCUMENT RESUME

ED 346 116 TM 018 357

TITLE	New York State Teacher Certification Examinations: Symposium Presented at the Annual Conference of the National Council on Measurement in Education (San Francisco, California, April 21-23, 1992).
INSTITUTION	National Evaluation Systems, Inc., Amherst, Mass.
PUB DATE	Apr 92
NOTE	41p.
PUB TYPE	Collected Works - Conference Proceedings (021)
EDRS PRICE	MF01/PC02 Plus Postage.
DESCRIPTORS	Conference Papers; Educational Assessment; Educational Policy; Elementary School Teachers; Elementary Secondary Education; Higher Education; Knowledge Level; *Licensing Examinations (Professions); Secondary School Teachers; *State Programs; *Teacher Certification; Teacher Education; Teaching (Occupation); Teaching Skills; Test Construction; *Testing Programs
IDENTIFIERS	*New York State Teacher Certification Examinations; *Subject Content Knowledge; Teacher Competency Testing

ABSTRACT

This collection of five symposium papers explores the initial design and development of the New York State Teacher Certification Examinations. The papers refer to work in progress to assess knowledge and skills associated with subject-matter knowledge and knowledge and skills in instruction and the profession of teaching. The following papers are included: (1) "State Certification Policy Issues and the Mandate for Innovative Teacher Certification Tests for New York State" (C. C. Mackey, Jr., and G. L. Freeborne); (2) "Testing Program Design and Development Procedures" (W. P. Gorth, R. G. Allan, and P. M. Nassif); (3) "Assuring Educated Educators through the New York State Liberal Arts and Sciences Test" (J. R. Silvestro and S. E. Lake); (4) "The New Standard for Subject-Matter Knowledge Assessment: The New York State Teacher Certification Examinations" (J. W. Clayton and C. C. Mackey, Jr.); and (5) "New York State Assessment of Teaching Skills: A Multiple Method Assessment of Teacher Competency" (S. Elliot and E. Hunsberger). A 22-item list of references is included. (SLD)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED346116

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

☑ This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

JOHN R. SILVESTRO

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

New York State Teacher Certification Examinations

Symposium presented
at the
Annual Conference
of the
National Council on Measurement in Education

SAN FRANCISCO 1992

TM019357

BEST COPY AVAILABLE

PREFACE

As part of the education reform movement, states are taking a closer look at the overall quality of the teaching profession. Most states have accepted the assumption that one key component of educational improvement is to establish meaningful and rigorous assessment programs as part of the teacher certification process. One state that has shown significant leadership in this regard is New York State. Recent amendments to New York State policies on the certification of elementary and secondary school teachers, have resulted in the creation of a comprehensive system of teacher certification assessment.

The New York State Teacher Certification Examinations (NYSTCE) represent a multifaceted assessment system. Included in this system are instruments to assess knowledge and skills associated with the liberal arts and sciences, knowledge of the subject-matter field in the candidate's intended area of teaching certification, and knowledge and skills in the area of instruction and the profession of teaching. Each component part of the NYSTCE fits into an integrated whole, and provides for a type and level of assessment that is appropriate to the level of training and expected expertise of developing professional educators.

This collection of papers, presented at a symposium offered at the 1992 annual conference of the National Conference of Measurement in Education, explores the initial design and development of the NYSTCE. As such, these papers refer to work that is currently in progress. Over the next 18 to 24 months, this work will take shape, no doubt with constructive revisions based upon the feedback and input of thousands of New York State educators, teacher educators, and members of professional teacher organizations.

The development of the NYSTCE has been an exhilarating assessment experience to date. The papers in this symposium provide examples of a number of dramatic innovations in the assessment of teacher knowledge and skills, including some measures of performance skills that break new ground in teacher assessment. It is hoped that this collection of papers will be the first in a series of reports that will be presented in subsequent years on what is likely to become a landmark program in teacher certification assessment.

John R. Silvestro
National Evaluation Systems
Symposium Organizer

CONTENTS

	Page
The New York State Teacher Certification Examinations (NYSTCE) **State Certification Policy Issues and the Mandate for Innovative Teacher Certification Tests for New York State** *Charles C. Mackey, Jr., Gerald L. Freeborne*	1
The New York State Teacher Certification Examinations (NYSTCE) **Testing Program Design and Development Procedures** *William P. Gorth, Richard G. Allan, Paula M. Nassif*	7
The New York State Teacher Certification Examinations (NYSTCE) **Assuring Educated Educators Through the Liberal Arts and Sciences Test** *John R. Silvestro, Sanford E. Lake*	15
The New York State Teacher Certification Examinations (NYSTCE) **The New Standard for Subject-Matter Knowledge Assessment** *Jeanne W. Clayton, Charles C. Mackey, Jr.*	21
The New York State Teacher Certification Examinations (NYSTCE) **New York State Assessment of Teaching Skills: A Multiple Method Assessment of Teacher Competency** *Scott M. Elliot, Edith Hunsberger*	26
REFERENCES	36

THE NEW YORK STATE TEACHER CERTIFICATION EXAMINATIONS (NYSTCE)

State Certification Policy Issues and the Mandate for Innovative Teacher Certification Tests for New York State

Charles C. Mackey, Jr.
Gerald L. Freeborne

New York State Education Department
Nelson A. Rockefeller Empire State Plaza
Albany, New York 12234

April 1992

Paper presented at the 1992 annual meeting of the National Council on Measurement in Education

Copyright © 1992 by the New York State Education Department.

THE NEW YORK STATE TEACHER CERTIFICATION EXAMINATIONS (NYSTCE)

State Certification Policy Issues and the Mandate for Innovative Teacher Certification Tests for New York State

Public education is, as we know, a state responsibility. The tenth amendment to the Federal Constitution gives the states those powers neither reserved for the Federal Government nor specifically denied the states. Each state derives from that amendment the authority to determine the conditions under which the schools will be established, the qualifications for the individuals who will staff the schools, and the curriculum of the schools.

Traditionally, state education agencies have been given broad functional responsibilities by their respective state statutes. The primary functions of state education departments are the following: leadership, regulation, and operation. The regulatory function includes, among others, the responsibility for approving or registering teacher education programs, applying sanctions, supervising the schools and distributing public funds under the public protection clause of our respective state constitutions. This function serves to help assure an educated public and to protect the lives and health of the state's youth. As a result, the state has the responsibility to establish minimal standards that are universally applicable. In large part, the regulatory function has dealt with the responsibility for public resources particularly as they have been expended on the employment of professional staffs in state public schools. In this regard, state regulations are designed to guarantee wise and economic use of educational funds and to ensure effective management of the educational enterprise.

Over the last decade, the Board of Regents has taken significant steps to improve the quality of education in New York State. Because student competence is the paramount goal of the education system, action has been targeted on strengthening pupil performance in elementary, middle, and secondary education.

While evidence is available to confirm that positive change is occurring, a number of social and economic developments sustain the need for continued measures to improve the effectiveness of our schools. The dropout problem and youth unemployment have become major national issues. The population as a whole has decreased in New York State, but the percentage of minority and low income families has increased these past ten years. New York State has felt the effects of a changing economy resulting in the loss of approximately 700,000 jobs in the last decade. The "technical revolution" has changed the nature of the job market. There is a widening gap between the advanced skills required by high technology industry and the skill levels typically associated with academic competencies.

Recognizing the complexity of the problem, the Regents and the State Education Department have taken a comprehensive approach toward the strengthening of pupil performance which involves a number of interconnected steps. This approach focuses on the three major components of the education system: the students, the educational programs, and teaching. The students and educational programs have been the focus of the Regents and Department's efforts to date, vis., Action Plan to Improve Elementary and Secondary Education Results in New York State. Now the focus is shifting to teachers and teaching.

The issuance of teachers' certificates has been a commonly accepted regulatory responsibility of the State Education Department. In New York State, the term teacher is used in the generic sense, meaning any person for whom a certificate is required in order to be legally employed in a public school. While specific standards vary markedly among the certifying agencies, it has been generally agreed that a totally satisfactory teacher certification procedure has yet to be devised.

New York State's current licensing system is an attempt to balance the need for more thorough teacher preparation with the realities of supply and demand. A major consideration in this equation is the pragmatic need for most teacher candidates to earn a living as they proceed through the teacher certification process. The assessment system which supports this process represents the recommendation of the 1977 Task Force on Teacher Education and Certification as articulated in statute and regulation. In the absence of sufficient resources to implement a testing process, an existing but recently amended nationally-based test battery was validated for use in New York State. These multiple-choice tests, which focus on communication skills, general knowledge, and professional knowledge, have long been viewed as interim assessment procedures pending the availability of resources or an innovative arrangement which would bring to this State a rigorous yet valid assessment system carefully designed to bring New York State teachers into the 21st century.

Early in the last decade a myriad of national reports were authored which, in sum or in part, focused on strengthening teachers and teaching. Among these were the Holmes Report; A Nation at Risk: The Imperative for Educational Reform; A Nation Prepared: Teachers for the 21st Century; and A Blueprint for Learning and Teaching (the last a report on the New York State Education Commissioner's Task Force on the Teaching Profession). Based on analytical studies of these reports, reflections on the needs of this State's educational community, a series of Board of Regents proposals to advance the recruitment, preparation and retention of a quality teaching force in New York State and a planned series of forums among all segments of the New York State "family", the Board of Regents on May 19, 1989, adopted new certification requirements which impact certificate titles representing 80% of the State's teaching force. These amended regulations carry the endorsement of the organized profession, the major constituency of the Commissioner's Task Force on the Teaching Profession; the Confederated Organizations for Teacher Education (an umbrella organization of teacher educators and institutions of higher education); and the Teacher Education, Certification and Practice Board (an advisory board to the State Board of Regents, the Commissioner and the Department).

Adopted effective September 2, 1993, the new requirements for the certification of teachers of common branch subjects in pre-kindergarten and grades 1-6 of the State's public elementary schools and for teachers of the academic subjects [English, a language other than English, mathematics, a science (biology, chemistry, physics, and earth science) and social studies in grades 7-12] of the State's public secondary schools will ensure to the public, said Commissioner Sobol, "new teachers will know their subject and how to teach it effectively to their students".

The new statewide teacher certification system will guarantee that before teachers are allowed to work independently in a classroom they will be thoroughly prepared for that important responsibility. The three general goals of this preparation are: to ensure a broad base of knowledge, to provide a strong foundation in pedagogy, and to ensure mastery of specific subject areas.

BEST COPY AVAILABLE

In each instance, the earlier cited certificate titles may be achieved by one of three means:

1. completion of a program registered (approved) by the Department for the preparation of teachers for that specific certificate title; or

2. the interstate agreement on qualifications of educational personnel, i.e., the interstate certification compact, or

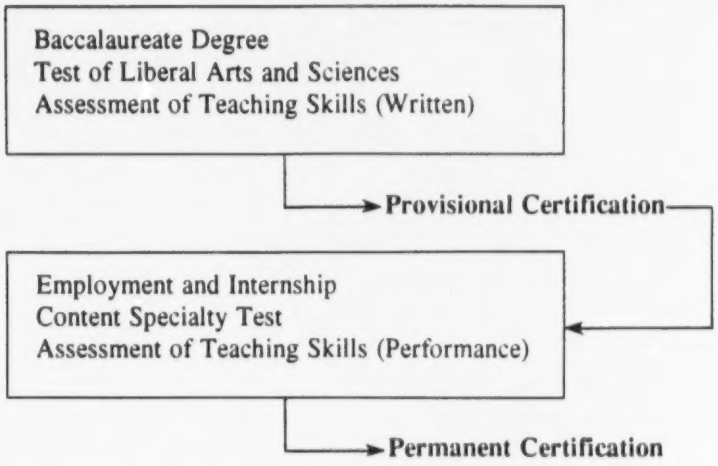
3. alternate completion of requirements, i.e., satisfaction of specified academic and experiential requirements through a review process.

With regard to the last, the Department will establish college faculty committees on academic and professional preparation to develop standards by which to evaluate concentrations in professional education, the liberal arts and sciences, and/or the certificate title specialization of individuals who apply for a State certificate but have not completed a concentration at an institution of higher education registered by the Department or approved by the Commissioner of Education. Provisions will also be made to assess teaching experience gained in nontraditional ways. The Regulations will continue to provide for a provisional certificate valid for five years and a permanent certificate valid for the life of the holder, unless annulled for cause.

For the provisional form of the two certificate titles identified earlier, common features are as follows: an earned baccalaureate degree with a concentration in one of the liberal arts and sciences, college-level study in a language other than English and qualifying scores on a new test of the liberal arts and sciences and on a new written assessment of teaching skills. For the permanent form of these certificates, an earned master's degree functionally related to the field of teaching service; a satisfactory one year supervised internship in a public or nonpublic school pursuant to regulations yet to be adopted by the Regents, and qualifying scores on an examination in the area of the teaching certificate title and on a performance assessment of teaching skills.

Within the context of the approved program, students preparing for careers as teachers of the common branch subjects in pre-kindergarten and grades 1-6 and of the academic subjects in grades 7-12 will be prepared to create a productive learning environment; monitor and assess student learning; address the special developmental and educational needs of students at all levels covered by the certificate; and work effectively with students from minority cultures, students from homes where English is not spoken, students with handicapping conditions, students of both sexes, and gifted and talented students. For students completing an approved program leading to the preK-6 certificate, college-level work in English, mathematics, science and social studies as well as special training in the teaching of reading and college-supervised field experiences in both the lower (preK-3) and upper (4-6) elementary-level grades will be included. For students completing approved programs leading to certification as teachers of academic subjects in grades 7-12, the curriculum will show evidence of a concentration of study in one of the liberal arts and sciences appropriate to the area of the teaching certificate and college-supervised field experiences at both the middle level (7-9) and high school grades (10-12).

The following schematic indicates the progression of this preparation which will culminate in the issuance of a permanent certificate:



The capstone of the new program for awarding teaching certificates is the well-defined assessment system to support decisions to award to teacher candidates provisional and permanent credentials in 20 certificate titles. The strategy of program development entails collaboration among State Education Department staff, practitioners in the State's public schools, teacher educators in the New York State institutions of higher education, and staff of National Evaluation Systems to produce a system that is technically, legally defensible and professionally appropriate. A teacher certification testing program will achieve acceptance and respect only insofar as its design is perceived to be based on an understanding of both the teaching profession and relevant psychometric standards and its implementation is shown to involve substantive input from all constituencies affected by the program.

In formatting this assessment system, the State Board of Regents and Commissioner of Education emphasized the need for one designed by and for New York State teachers. As such, it was necessary that this be a custom designed program, not an "off the shelf" test, correlated with New York State teacher education programs and public school practice. Because the assessment system will be designed and developed as a new program from beginning to end, the Department will retain control of the program. Further, the program will not remain static. As state standards and regulations are updated and new educational policies and content directions emerge, New York State will be able to review and revise the assessment program to maintain its currency.

As proposed in The Regulations of the Commissioner of Education, the assessment system for provisional and permanent certification must be integrated to ensure that the tests are reflective of the differential knowledge and skill needs of beginning and continuing public school teachers. Consistent with this goal was the State's desire to assess performance on a test of liberal arts and sciences that would legitimately reflect the thinking skills acquired in the preprofessional component of the preservice program of preparation. This test of liberal arts and science will cover historical and social scientific consciousness, scientific and mathematical processes, and artistic expression and the humanities. Candidates taking the examination are to demonstrate ability in four categories: modes of inquire/conceptual content/principles; multicultural awareness; communication skills; and critical thinking skills.

For the provisional or entry-level certificate, the candidate must also complete the first part (written) of an assessment of teaching skills. This test, which will address the theoretical aspects of the pedagogical preparation, will be a written examination consisting of both an exercise requiring an extended written response and a series of multiple-choice items.

Also, consistent with this goal of differential knowledge and skill, the State is committed to developing tests of content knowledge in the certificate title (e.g., early childhood or elementary education, English, mathematics) that would be reflective of the knowledge mastered by those completing a concentration in the field of certification, not just the knowledge required of a prospective educator in the field. The content specialty test, however, will not assess pedagogical skill, since pedagogy is covered by the Assessment of Teaching Skills, both in the written and the performance portions.

The second, or performance part of the Assessment of Teaching Skills is a requirement for permanent certification. Here, preliminary plans indicate that the candidate may be asked to submit a videotape of a class that he or she has taught with a completed form explaining the background, goals, and context for the lesson and providing brief commentary on the success of the session. The candidate may be asked to show cooperative learning, one-to-one instruction, and the more traditional whole-class instruction. The tapes would be scored by New York State teachers trained for conducting the assessments. Separate examinations of teaching skills will be developed for the elementary and for the middle and secondary levels. Passing this video performance assessment is predicated on the candidate successfully completing the written component for the provisional certificate.

Finally, the Regents and Commissioner determined a need to balance traditional testing formats with alternative testing formats, in order to provide for more realistic assessments of what new teachers are actually expected to do in their respective classrooms. National Evaluation Systems is currently implementing an array of alternative assessment strategies in several of its programs; these will be incorporated as their appropriateness to the objectives of this testing program is established. Among these alternative methodologies are the following: assessment of listening comprehension through tape recorded passages and dialogues, items for which an examinee-created short answer is required, assignments that elicit examinee writing samples or other samples of performance, items for which examinees must select more than one correct answer, problems that focus on issues in classroom or student management, written simulations of important situations that occur in real-life contexts, and videotaping classroom teaching performance.

While many alternative assessment approaches have been used in practical situations, the issue of appropriateness if, of course, the key issue in deciding on any assessment strategy. Is the contemplated method appropriate for the situation rather than being merely new and interesting? In this regard, the particular issues to consider include validity, reliability, freedom from bias, feasibility, and cost-effectiveness.

In summary, we believe the standards which New York State teachers should know and be able to demonstrate are rigorous and will represent a level and scope of professionalism unmatched in any other licensing jurisdiction.

THE NEW YORK STATE TEACHER CERTIFICATION EXAMINATIONS (NYSTCE)

Testing Program Design and Development Procedures

William P. Gorth
Richard G. Allan
Paula M. Nassif

National Evaluation Systems, Inc.
30 Gatehouse Road
Amherst, Massachusetts 01004

April 1992

Paper presented at the 1992 annual meeting of the National Council on Measurement in Education

Copyright © 1992 by National Evaluation Systems, Inc. (NES®). All rights reserved.

"NES®," "National Evaluation Systems, Inc. ™," and the "NES®" logo are trademarks of National Evaluation Systems, Inc.

THE NEW YORK STATE TEACHER CERTIFICATION EXAMINATIONS (NYSTCE)

Testing Program Design and Development Procedures

BACKGROUND

Program Purpose

Among the purposes for the New York State Teacher Certification Examinations program is the desire to develop an integrated system of meaningful and authentic assessment for teaching certification candidates. Toward this end, all three tests will require candidates to demonstrate the knowledge and skills that would be expected of educators in the daily conduct of their teaching assignments. The assessments will form one component of a teacher certification testing system designed to help ensure that New York State educators have the knowledge and skills needed to perform the ever more demanding job of the classroom teacher. Key among any plans for improving the quality of education is the role of the classroom teacher. While there are a large number of factors that can affect the schools, the classroom teacher is the central figure in any educational reform and improvement plan. The new testing program will be custom developed to meet unique New York State needs and specifications and is being undertaken with the extensive involvement of New York State educators.

Each component in the assessment program will be based upon a substantive definition of knowledge and skills required of an entry-level educator in New York State public schools. The objectives defining each assessment will be based in part on recent research on teaching, New York State policy and curriculum documents and the input of practicing classroom teachers and teacher educators, particularly in New York State.

Structure of the Program

The structure of the New York State Teacher Certification Examinations program includes three primary assessment areas:

1. **The Liberal Arts and Sciences Test**. This test will be designed to assess a broad base of knowledge and skills across the disciplines (historical and social scientific, mathematics and science, arts and humanities). Assessments will be designed to measure candidates' abilities in four categories: 1) modes of inquiry/concepts and principles, 2) multicultural awareness and sensitivity, 3) communication skills, and 4) critical thinking skills. This test will be required for initial, provisional certification.

2. **Subject-Matter Tests**. These tests will be designed to assess the content knowledge of candidates for specific certificates in a number of areas. Types of content to be assessed include factual content, modes of inquiry, concepts and principles, applications of content, relationships of content to other areas both inside and outside the discipline, and problem solving. These tests will be required for permanent certification.

3. **The Assessment of Teaching Skills**. This assessment is composed of two parts, a written assessment and a performance assessment. The assessments will evaluate the pedagogical knowledge and skills of candidates. Areas of knowledge and skill to be assessed include development in learning, instructional planning and curriculum development, delivery of instruction, classroom management, assessment and evaluation, and statutes, regulations, and governance. The written portion of this assessment will be required for provisional certification and the performance component will be required for permanent certification.

New York State Educator Involvement

Advisory Committees. The involvement of New York State educators is central to the development process for the New York State Teacher Certification Examinations. The following New York State educators will participate in the test development process.

Bias Review Committee. The Bias Review Committee established by the New York State Education Department is composed of minority New York State educators who review the test materials for sensitivity and fairness. The committee has also been involved in the review and updating of National Evaluation Systems' publication *Bias Issues in Test Development*.

Content Advisory Committees. The Content Advisory Committees of New York State educators established by the New York State Education Department include public school educators and teacher educators. The committees review draft test materials according to criteria established for the program.

Job Analysis Survey Participants. New York State educators are surveyed to gather additional New York State educator input on the proposed test objectives.

Alternative Assessment Formats

The examinations will use performance-based assessment formats that extend beyond the traditional multiple-choice item formats. The particular applications of each format will be discussed in more detail in the remaining papers in this symposium.

Written Assignments. Both the Liberal Arts and Sciences and Assessment of Teaching Skills assessment will require examinees to provide a written response to a given prompt. For Liberal Arts and Sciences, examinee responses will be evaluated in relation to their ability to communicate effectively in writing, their ability to demonstrate critical thinking skills and the skills expected of a liberally educated individual. For the assessment of teaching skills, examinees may be asked to analyze an educational issue, solve an educational problem presented or produce a work sample.

Speaking Assessments. The foreign language assessments will require examinees to demonstrate oral proficiency in a performance assessment. Given a situation, question, or prompt, examinees will be asked to give a spoken response, which will be recorded on audiotape and evaluated by New York State foreign language educators.

Listening Assessments. The foreign language assessment will also require examinees to listen to audio stimuli and respond in writing or to multiple-choice items.

BEST COPY AVAILABLE

Videobased Performance Assessment. The Assessment of Teaching Skills may include a videotaped performance assessment. Candidates would provide a self-made video tape of their classroom performance along with written documentation of the objectives of the lesson and class characteristics. The videotaped performance would be scored by New York State educators based on criteria related to classroom teaching skills.

Support for Examinees and Teacher Preparation Institutions

Procedures will be established that assist individuals in test preparation and in the remediation of knowledge and skill deficiencies if the first test attempt is unsuccessful. Information will be provided to the colleges and universities to enable them to assist faculty and students in strengthening prospective examinee performance on the tests. The objectives defining each test area will be disseminated to all teacher education programs in the state.

EXAMINATION DEVELOPMENT PROCESS

Development activities for the examinations began in 1991 and are continuing. The development activities are summarized below.

Defining the Content Knowledge and Skills to be Measured. For each test, a framework organizing the field into subareas of content, each structured and described by objectives, was developed to serve as an explicit foundation for test content. Each objective is further defined by focus statements. The frameworks will not only guide test development but will also assist candidates in their preparation for the examinations. The frameworks were developed utilizing New York State legislation and regulations, New York State public school curriculum materials, New York State college and university course catalogues, syllabi, and textbooks; and other curriculum materials from New York State higher education institutions. The frameworks were reviewed by New York State curriculum specialists and the NYSED, and were reviewed and finalized by the Bias Review Committee and the Content Advisory Committee.

Develop Assessment Specifications. Assessment specifications were written to provide guidance to item writers on what content should be covered on the test and how questions should be written. The assessment specifications include sample items to illustrate the style and format of the items to be developed. The Bias Review Committee and Content Advisory Committee reviewed and finalized the assessment specifications for each field.

Job Analysis Survey. The approved framework of objectives for each test will be included in a job analysis survey to establish that the objectives are related to the job of an educator in New York State public schools. Practicing New York State public school teachers in each field will be asked to rate the importance of each objective to the job of a New York State teacher in that field. In a separate survey, New York State teacher educators will also be asked to rate the importance of each objective to the job of a New York State teacher in that field. The results of the survey will be presented to the Bias Review Committee and Content Advisory Committee.

Develop Assessment Materials. Assessments will be developed to correspond to the approved objectives and assessment specifications, and will be field tested. A number of different assessment methods and formats may be used, as described earlier.

Conduct Assessment Review Meetings. Draft assessments will be reviewed and revised by both the Bias Review Committee and the Content Advisory Committee for each field.

Conduct Item Revalidation and Standard Setting. The Content Advisory Committee will review each item again for validity and to provide judgments that will be used by the NYSED in setting a passing score for each test.

Validity

Validity is a primary concern in any measurement effort. Validity has been defined as "an integrated evaluative judgment of the degree to which empirical evidence and theoretical rationales support the adequacy and appropriateness of inferences and actions based on test scores or other modes of assessment" (Messick, 1989). Making this judgment usually involves a process of accumulating evidence from a variety of sources in support of the inferences made (American Psychological Association [APA] Standards, 1985).

In employment testing, the primary focus is on establishing that the assessment instrument measures the aspects of job performance that it purports to measure. Obtaining evidence in support of this construct may take several forms. Documentary evidence, expert judgment, and empirical data can be brought to bear to support the determination of whether individuals are competent to perform the job.

NYSTCE Validation Approach

One of the primary concerns in designing the NYSTCE program was test validity. There are several key features to the validation approach designed for the NYSTCE program: establishing validation procedures at each step in the test development process, using multiple sources of validity evidence, maintaining consistency with New York State public policy, providing empirical links to New York State educational practice, and consulting advisory committees. These key features are described below.

Validation at Each Test Development Step

The validity literature (cf. Cronbach, 1971; APA *Standards*, 1985; Messick, 1989) emphasizes that validation is a process of accumulating evidence in support of a test rather than a single event occurring at one point in time. While many test development efforts rely on a single validation procedure (typically post hoc), the decision was made at the outset of the NYSTCE program to include validation procedures at each major step in the test development process. Steps to ensure validity were taken in defining the test domain, conducting the job analysis, and constructing items.

Multiple Sources of Evidence

Multiple sources of validity evidence were obtained throughout program development. The major sources were New York State public policy and New York State educational practice.

New York State Public Policy. As described by Freeborne and Mackey (1992) in this set of papers, New York State public policy statements were actively incorporated into the test development process in order to ensure that the test would be valid for the purpose for which it was intended. These policy statements provided the foundation for test development activities.

New York State Education Practice. Both the Equal Employment Opportunity Commission (EEOC) guidelines (1978) and the APA *Standards* (1985) address the need to establish an empirical link between the content of the job and the domain of the test. To ensure that the NYSTCE reflected educational practice in New York State, a systematic job analysis of practicing public school teachers in each area in which tests were to be developed was conducted. To provide further information about educational practice in the state, teacher educators responsible for preparing New York State teachers were surveyed to verify the job relatedness of the proposed test content.

Advisory Committees (Expert Judgment). As a further measure to ensure that the NYSTCE reflected educational practice in New York State, the NYSED convened a Content Advisory Committee composed of New York State educators in each teaching field for which tests were being developed. The committee was responsible for reviewing and making judgments about program materials at several points in the test development process. In addition to this advisory committee, a second committee composed of New York State educators familiar with bias issues, the Bias Review Committee, convened to review materials for potential bias.

Reliability

Reliability, which relates to the extent to which a measure consistently produces the same result under similar conditions (Nunnally, 1978), is a major concern with teacher assessment. Traditionally, reliability has been thought of as the internal consistency of a test or the stability of test scores across repeated administrations (test-retest reliability) and parallel forms of the test (equivalence). More recently, particularly in the area of certification, test developers have begun to examine reliability in terms of the dependability of classification decisions (e.g., pass-fail).

Classification Decisions. A number of writers (Livingston, 1972; Huynh, 1976; Berk, 1980) have suggested that the reliability of tests in situations where a dichotomous decision is made (e.g., pass-fail) should be evaluated with respect to the consistency with which those decisions are made. Estimates of the reliability of classification decisions fall primarily into two categories: (1) reliability indices that provide an estimate of the consistency of mastery and nonmastery decisions based on a threshold loss function and (2) indices that focus on the consistency of criterion-referenced scores across the distribution, based on the squared error loss function. Essentially, the focus of the first approach is on the reliability of the decision made, while the focus of the second approach is on the reliability of all scores obtained on a criterion-referenced test. These approaches are reviewed in detail by Berk (1980) and others.

Multiple-Choice Reliability. The reliability of the classification decisions for the multiple-choice items for the NYSTCE will be estimated using the methods suggested by Huynh (1976) and Livingston (1972).

Performance Assessment Reliability. While much attention has been focused on the reliability of criterion-referenced tests comprising dichotomously scored items, far less attention has been given to applying similar estimation techniques to polytomously scored items such as those included on the performance-based components of the NYSTCE. The reliability of the written assignment section will be computed using a generalizability theory approach. This technique will provide an estimate of the stability of the ability estimates provided by the performance assessment components.

Eliminating Bias

Ongoing efforts will be made during NYSTCE program development to eliminate potential bias against groups of examinees on the basis of irrelevant factors or characteristics. These efforts will focus on a combination of professional judgments about the appropriateness and freedom from bias of program materials and the gathering and interpretation of statistical information about differential item functioning. Professional judgments will be applied at major review points throughout the process. Statistical data will be generated from test administration data collected.

This combination of professional judgment and statistical analysis has been recommended as a desirable practice within the testing profession (Shepard, 1982). Reliance on both professional reviews and statistical information is considered appropriate for several reasons.

It has been suggested that the construct of bias is multidimensional (Berk, 1982) and that judgmental reviews and statistical methods of bias detection should complement each other. According to this view, each method may contribute its own separate strengths to the analysis of potential bias. Statistical analysis is strongest in detecting test items that produce larger than expected group differences in performance (Sandoval & Miille, 1980; Plake, 1980). Professional reviewers may focus on aspects of the bias construct (e.g., stereotyping) that it is highly desirable to eliminate from test materials but that might have either no negative effect on examinee performance or no locally detectable effect but only a more subtle, cumulative effect over an entire test or set of tests (Tittle, 1982).

The application of professional judgment is best conceptualized not as a separate activity from the gathering and interpretation of statistical information, but rather as an important accompaniment. For these reasons, it was considered most appropriate for the NYSTCE to build into the program design for bias prevention the complementary application of a judgmental component and a statistical analysis component.

JUDGMENTAL REVIEWS

Professional judgments intended to detect potential bias in the NYSTCE objectives assessment specifications, multiple-choice test items, and written assignments were gathered throughout the development of the program. For purposes of the program, an inclusive definition of the concept of potential bias was used, addressing not only the *presence* of potentially biasing content, language, and ideas but also the *absence* of positive portrayals of diversity (Tittle, 1982).

Review Criteria

The review criteria that were applied in judgmental reviews of the list of draft objectives and assessment specifications for the NYSTCE program, which are similar to the criteria used for other test materials, convey the inclusive definition of the bias concept that was used throughout the program.

Content. Does any element of the list [of objectives or assessment specifications] contain content that disadvantages a person because of his or her gender, race, nationality, ethnicity, religion, age, handicapping condition, or cultural, economic, or geographic background?

Language. Does the language used to describe any element of the list disadvantage a person because of his or her gender, race, nationality, ethnicity, religion, age, handicapping condition, or cultural, economic, or geographic background?

Offense. Is any element of the list presented in such a way as to offend a person because of his or her gender, race, nationality, ethnicity, religion, age, handicapping condition, or cultural, economic, or geographic background?

Stereotypes. Does any element of the list contain language or content that reflects a stereotypical view of a group based on gender, race, nationality, ethnicity, religion, age, handicapping condition, or cultural, economic, or geographic background?

Diversity. Does the list permit appropriate inclusion of content that reflects the diversity of the New York State population?

This approach to review criteria was designed to elicit more than attempts by reviewers to detect and remedy potential bias in the assessment materials. The goal was to assume an affirmative posture relative to the elimination of potential bias by ensuring that positive aspects of diversity were present in the materials as well.

Groups Reviewing Materials for Potential Bias

During the development of the program, several groups of individuals will take part in reviews of test materials specifically for bias issues. Three review groups in particular will have a substantial role in this effort.

NES Equity Advisory Board. The NES Equity Advisory Board is a standing group of educators who have experience in dealing with issues relating to potential bias in educational programs and materials. For the NYSTCE, members of the Equity Advisory Board reviewed draft objectives and assessment specifications, and will review draft items.

NYSTCE Bias Review Committee. The members of the Bias Review Committee were selected by the NYSED to review materials for potential bias. This committee reviewed objectives and assessment specifications, and will review draft items.

NYSTCE Content Advisory Committees. The Content Advisory Committees reviewed draft objectives and assessment specifications, and will review draft items. These committees independently review materials for bias and review the recommendations of the Bias Review Committee, whose review meetings will preceded theirs.

Summary

The New York State Teacher Certification Examinations will provide a comprehensive, three-tiered assessment of teachers, using a wide variety of assessment methods. The test development process will focus on extensive involvement of New York State educators and a continuing commitment to fairness.

THE NEW YORK STATE TEACHER CERTIFICATION EXAMINATIONS (NYSTCE)

Assuring Educated Educators Through the New York State Liberal Arts and Sciences Test

John R. Silvestro
National Evaluation Systems
30 Gatehouse Road
Amherst, Massachusetts 01004

Sanford E. Lake
New York State Education Department
Nelson A. Rockefeller Empire State Plaza
Albany, New York 12234

April 1992

Paper presented at the 1992 annual meeting of the National Council on Measurement in Education

Copyright © 1992 by National Evaluation Systems, Inc. (NES®). All rights reserved.

"NES®," "National Evaluation Systems, Inc. ™," and the "NES®" logo are trademarks of National Evaluation Systems, Inc.

THE NEW YORK STATE TEACHER CERTIFICATION EXAMINATIONS (NYSTCE)

Assuring Educated Educators Through the New York State Liberal Arts and Sciences Test

Introduction

One of the more frequent concerns voiced about the quality of American educators by those who have been involved in studying educational reform, has been that a significant segment of the American teaching force is unable to demonstrate many of the intellectual skills that are characteristic of the "well-educated person". This is particularly notable in the relative ability of teachers to integrate the liberal arts and sciences into their classroom instruction, and the paucity of such transdisciplinary perspectives and practices. It is essential that teachers understand not only their respective area(s) of concentration and the educational needs of their students but, also, that they recognize the constructs upon which curricula are predicated and the importance of addressing classroom subjects from multiple perspectives within multiple theoretical constructs. There are, no doubt, many reasons for this inability of all too many teachers to integrate the liberal arts and sciences into their practice. Some critics point to the traditional teacher training programs offered by institutions of higher education as the primary reason why teachers show deficiencies in the liberal arts and sciences. They note that some schools of education have sacrificed intellectual breadth, typically acquired through a range of arts and sciences requirements and electives, for an overemphasis on professional education courses. Other critics do not cast blame on the schools of education but rather point to a decline in the general liberal education skills of many college students, even those majoring in traditional liberal arts fields of study. These latter critics blame the pandemic dilution of the core liberal arts college curriculum requirements that took place in the 1970s as the primary cause of liberal arts and sciences deficiencies among college graduates in general, not only teacher education graduates.

Regardless of the cause of the decline and fall of liberal education, there is growing concern by both universities in general,and more specifically by teacher education colleges, that those who profess to be educators must also be educated in the broadest sense of the term. As a result of the efforts of the Holmes Group (1986) and the principles presented in *A Nation Prepared: Teachers for the 21st Century* (Carnegie Forum on Education and the Economy, 1986) and other contemporary examinations of American education, there has been a resurgence of interest in strengthening the general liberal educational skills of prospective teachers. New York State has been at the forefront of the effort to ensure that its candidates for teaching certificates demonstrate satisfactory levels of mastery of the intellectual skills that are the hallmark of an educated person. Sections 80.15 and 80.16 of *Title 8 of the Codes, Rules and Regulations of the State of New York, Regulations of the Commissioner of Education*, effective September 2, 1993, mandates for the provisional and permanent common branch and secondary academic teaching certificates in grades PreK-6 and 7-12, respectively, requirements inclusive of satisfactory performance on a test of the liberal arts and sciences and a concentration in at least one of the liberal arts and sciences.

BEST COPY AVAILABLE

Test Design

With the establishment of new certification requirements that are designed to strengthen and broaden the preparation, training and assessment of candidates for provisional and permanent certificates as teachers in the public elementary and secondary schools, the New York State Education Department contracted with National Evaluation Systems (NES®) to develop a test of the liberal arts and sciences. From the initial stages of test development for the Liberal Arts and Sciences Test (LAST), the New York State Education Department has made it very clear to NES what the overall philosophy of the LAST should be. The design of the LAST is predicated on the belief that a well-educated teacher, in addition to being exposed to the content knowledge presented in liberal arts and sciences courses, should have also acquired generalizable thinking skills that help that teacher see the connections among ideas and facts and to be able to apply the information and methods of inquiry of the major knowledge domains to a variety of situations. For example, a social studies teacher would benefit from having knowledge from not only the social studies domain, but also from the humanities, mathematics, natural sciences, the arts, and communication. For a social studies teacher to effectively provide instruction on the European Renaissance, he or she would have to integrate knowledge from the diverse disciplines noted above and would have to be able to apply the critical thinking skills of these disciplines to help students learn and acquire meaningful knowledge about the Renaissance and the relative importance of this historical period in different contexts.

Given this basic orientation, test items on the LAST will be designed to assess one or more of the following cognitive skills:

- a general understanding of major themes, concepts, and developments in each area tested;
- an understanding of relationships among major themes and concepts in each area;
- an ability to apply major themes and concepts to real-world issues, problems, and examples related to each area;
- an ability to apply higher-order and critical thinking skills in contexts related to each area; and
- an ability to apply the fundamental methods , processes, and ways of thinking that characterize each area.

The skills noted above will be assessed across five content domains or test subareas as follows:

I. Scientific and Mathematical Processes;

II. Historical and Social Scientific Awareness;

III. Artistic Expression and the Humanities;

IV. Communication Skills; and

V. Written Analysis and Expression.

Examples of Representative LAST Assessment Materials

Across the five LAST content domains, there will be a total of 21 test objectives. All LAST test items will be written to match a specific test objective. To demonstrate the types and levels of content and skills assessed by the LAST, a sample test objective is presented as follows:

Understand the interrelatedness of historical, geographic, cultural, economic, political, and social issues and factors.

The focus of the assessment of this test objective would likely include the following dimensions:

* assessing the likely effects of human activities or trends (described in written or graphic form) on the local, regional, or global environment;
* assessing ways in which major transformations related to human work, thought, and belief (e.g., industrialization, the scientific revolution, the development and spread of Islam) have affected human society;
* inferring aspects of a society's social structure or group interactions based on information presented in an excerpt; and
* assessing the social or economic implications of political views presented in an excerpt.

Each test objective and the focus statements associated with each test objective, would be widely circulated to New York State teacher preparation institutions and to prospective examinees, to facilitate college planning and examinee preparation for the LAST.

Item Types and Formats

A variety of item types and formats will be included on the LAST, as indicated below.

Multiple-Choice Items. In the multiple-choice portion of the test, examinees are expected to demonstrate:

* knowledge of scientific and mathematical processes;
* historical and social science awareness;
* familiarity with artistic expression and the humanities; and
* skills of communication.

In some multiple-choice items, a stimulus is included to ensure that examinees are given adequate background information to display their knowledge, regardless of their academic major or area of specialization. Stimuli may be presented either singly or, for comparison purposes, in sets. Sets of stimuli may be presented in similar modes (e.g., two excerpts from historical texts) or in different modes (e.g., a textural excerpt, a graph, a table). More than one item may be related to a single stimulus or set of stimuli; items so grouped may come from different test objectives or subareas.

Written Assignment. The written assignment calls upon examinees to demonstrate the ability to communicate their thinking about a particular stimulus or set of stimuli, called a prompt, using the conventions of edited American English. Examinee responses will be approximately 300 to 600 words. The audience for the communication is to be a general audience of educated adults rather than an audience educated in the same discipline as the writer.

Prompts may be :

- one or more statements or excerpts relating to an issue of current concern;
- information presented in written, pictorial, graphic, or tabular form; or
- hypothetical situations involving ethical or value-related considerations.

Examinees' tasks may be to:

- analyze the information in the prompt;
- explain the information in the prompt;
- discuss ideas and issues related to the prompt;
- compare or contrast separate sources of information presented in the prompt; or
- persuade the reader of an opinion or course of action relating to issues presented in the prompt.

LAST Sensitivity to Diversity

The representation of cultural diversity in test stimuli and items is a major goal of the LAST. Stimuli and items will reflect the cultural diversity of the national and state populations. Historical and contemporary contributions and perspectives of women and minority groups will be addressed. In the selection of item stimuli, reading passages should:

- relate to a variety of time periods and regions;
- reflect gender and cultural diversity in content and authorship;
- be free of bias;
- represent individuals in positive roles as well as negative; and
- not appear to endorse points of view that could be considered offensive.

Summary

The New York State Liberal Arts and Sciences Test will, hopefully, play an integral role in helping to ensure that candidates for provisional teaching certificates possess the intellectual skills and associated knowledge that are reflective of a broadly educated person. More importantly, the LAST may help to ensure that certified teachers in New York State draw upon the knowledge and skills acquired from their liberal arts and sciences background, to enrich the overall quality of their instruction from the perspectives of content, student motivation, breadth of student cognitive skills elicited during instruction, and inclusion of elements of cultural diversity into the public school curriculum. The LAST is an ambitious assessment effort that will likely have a significant impact on future teacher certification testing in New York State and in other states that are concerned about preparing persons well-educated to provide instruction to elementary, middle and secondary school students.

THE NEW YORK STATE TEACHER CERTIFICATION EXAMINATIONS (NYSTCE)

The New Standard for Subject-Matter Knowledge Assessment: The New York State Teacher Certification Examinations

Jeanne W. Clayton
National Evaluation Systems, Inc.
30 Gatehouse Road
Amherst, Massachusetts 01004

Charles C. Mackey, Jr.
New York State Education Department
Nelson A. Rockefeller Empire State Plaza
Albany, New York 12234

April 1992

Paper presented at the 1992 annual meeting of the National Council on Measurement in Education
Copyright © 1992 by National Evaluation Systems, Inc. (NES®). All rights reserved.

"NES®," "National Evaluation Systems, Inc. ™," and the "NES®" logo are trademarks of National Evaluation Systems, Inc.

THE NEW YORK STATE TEACHER CERTIFICATION EXAMINATIONS (NYSTCE)

The New Standard for Subject-Matter Knowledge Assessment: The New York State Teacher Certification Examinations

BACKGROUND

Regulations of the New York State Commissioner of Education call for candidates seeking permanent certificates in elementary education or a secondary academic subject area to complete several requirements, including achieving a qualifying score on an examination in the area of the teaching certificate. There are presently nineteen certificate areas covered by these regulations: elementary education, English, mathematics, social studies, biology, chemistry, earth science, physics, and 11 languages other than English (Latin, Cantonese, French, German, Greek, Hebrew, Italian, Japanese, Mandarin, Russian, Spanish). New York State Teacher Certification Examinations (NYSTCE) are currently being developed for each of these nineteen content areas.

Goals of the Tests

As licensing tests for teachers seeking permanent certificates, these content specialty tests are intended to ensure that candidates have the content knowledge that is important to perform the job of an elementary or secondary academic teacher in New York State public schools. At the same time, the New York State Education Department (NYSED) is committed to developing subject matter tests that assess more than the instructional content that a teacher delivers in the classroom. The tests are designed to be reflective of the level of knowledge mastered by individuals completing a baccalaureate degree with a concentration in the title of the certificate to be awarded.

Targeting the test content at the level that is reflective of a college concentration is designed to ensure that candidates for teaching certificates have mastery of their subject area at a level that is important to effectively deliver the content. While this level of knowledge will generally not be taught directly to students, the knowledge may be called upon by teachers in designing instruction, answering student questions, evaluating student learning, and carrying out other responsibilities of a teacher.

For elementary education as well as the secondary academic areas, candidates are expected to have completed a college concentration in one of the liberal arts and sciences. The elementary education test assesses candidates' content knowledge in the domains of English language arts, mathematics, social studies, science, the fine arts, and health and fitness. While candidates for certification in elementary education cannot be expected to have acquired content knowledge at the level of a college concentration in each of these domains, candidates can be expected to have a level of knowledge reflective of individuals completing a concentration in one of the liberal arts and sciences.

The content specialty tests are not intended to measure pedagogical knowledge and skills. These aspects of teaching will be covered by the written and performance components of the Assessment of Teaching Skills (see Elliot and Hunsberger paper). The elementary education test and the tests in the secondary academic areas are designed to assess content knowledge that is both important to the teacher's job and reflective of the knowledge expected of a college graduate with a concentration in a related liberal arts and sciences area.

Sample Objectives

Following is a sample objective from the Social Studies test:

> Understand the major physical and cultural regions of the world, recognize geographic relationships within and between regions, and demonstrate an appreciation of the diverse perspectives of different human groups and cultures.

The focus statements, or examples, listed with this objective are as follows:

- demonstrating a familiarity with the major cultural groups associated with particular regions
- understanding the economic, environmental, and cultural factors contributing to demographic change
- analyzing the relationships between resource distribution, economic organization, and political authority
- analyzing ways in which political or cultural values have influenced developments in various regions or societies

Below is a sample objective, with its focus statements, from the English framework:

> Understand the historical, social, and cultural aspects of literature, including the ways in which literary works and movements both reflect and shape culture and history.
>
> For example:
>
> - applying knowledge of the characteristics and significance of mythology and folk literature
> - analyzing the expression of cultural values and ideas (e.g., regional, ethnic, historical) through literature
> - analyzing the role of given authors and works in influencing public opinion about and understanding of social issues (e.g., Harriet Beecher Stowe, Charles Dickens, Jonathan Swift, Toni Morrison, N. Scott Momaday, Tato Laviera)

Dimensions Assessed

The content specialty tests included in the New York State Teacher Certification Examinations are intended to measure more than recall of specific facts related to the content area. The tests will require understanding of concepts and principles and will require the ability to apply or analyze factual knowledge, concepts, and principles. The tests will require use of higher-order thinking skills and other specific skills related to the content area. The test items will not require recall of highly specific details, but will assess candidates' use of knowledge to understand significant themes, trends, and processes. Tests will include reading passages from a variety of sources as well as other stimuli such as graphs, tables, diagrams, and experimental data. Use of the stimuli will be required to demonstrate comprehension of relationships, analytic understanding, and the application of skills and knowledge to solve problems.

Item Formats

One of the goals of the New York State Education Department for the testing program is to balance traditional testing formats with alternative testing formats. This approach is designed to ensure that the testing instruments will be reliable, statistically sound, and appropriate as licensing tests, and at the same time, that they will include some direct measures of performance and other innovative item formats.

The content specialty tests should be viewed in the context of the entire testing program. The Liberal Arts and Sciences Test, which will call upon candidates to apply thinking skills to a variety of situations within major knowledge domains, will contain a variety of multiple-choice items and a written assignment that will assess candidates' communication skills. The two test components of the Assessment of Teaching Skills will assess candidates' pedagogical knowledge and skills through multiple-choice items, a written essay, and a direct assessment of teaching performance. The content specialty tests will assess candidates' content knowledge through a variety of item types appropriate for the content area.

The tests for the content areas will include a variety of multiple-choice items that will assess examinee knowledge, understanding, and skills in a reliable, sophisticated, and cost-effective manner. These items will include the use of written or graphic stimuli and will call on candidates to analyze information and to apply thinking skills to solve problems. Some stimuli will be clustered to a number of items and will call upon candidates to demonstrate understanding in a number of domains. Some items will present a number of conclusions, action statements, or other responses, of which more than one may be correct.

Each of the tests of languages other than English will require examinees to demonstrate proficiency through direct measures. In each of the 11 language tests, candidates will demonstrate reading, listening, speaking, and writing skills through direct measures. For the Latin test, candidates will read aloud passages of prose and poetry and will be required to demonstrate appropriate pronunciation and phrasing. Candidates will also listen to spoken recordings and transcribe the spoken passages or prose.

For the modern languages tests, examinees will demonstrate listening proficiency by listening to recorded messages. Candidates will be required to demonstrate comprehension of the message and understanding of the tone or point of view of the speaker or speakers. Candidates will also be required to derive information from oral messages of real-life situations, as well as inferring information that is implied but not stated.

Candidates for certification in each of the modern languages will also write a passage of several paragraphs and will be required to communicate a message effectively for a given audience and purpose. A prompt will be presented that may call on the candidate to write a letter requesting a recommendation for a job, or to describe the reason for a personal decision, or to communicate another specified message.

To demonstrate speaking proficiency, candidates will respond, on tape, to a prompt that may call for them to describe an event or action, to discuss advantages or disadvantages of an idea, or to respond to a hypothetical situation. The candidate will be expected to communicate a response effectively and to demonstrate command of a range of vocabulary and simple and complex language structures.

A Balanced Approach

The New York State Education Department has strived to take a balanced approach to test development. The structure of the testing program, the test development process, and the test design and item format decisions are all based on an approach of balancing needs of validity, reliability, feasibility, innovation, and cost-effectiveness. Each of the decisions that has been made for the program has been made thoughtfully and with the overall goals of the testing program in mind. The expectations and standards for the testing program have been maintained at a very high level, with the ultimate expectation that the testing program will improve the professionalism of teaching in New York State.

THE NEW YORK STATE TEACHER CERTIFICATION EXAMINATIONS (NYSTCE)

New York State Assessment of Teaching Skills: A Multiple Method Assessment of Teacher Competency [1]

Scott Elliot
National Evaluation Systems, Inc.
30 Gatehouse Road
Amherst, Massachusetts 01004

Edith Hunsberger
New York State Education Department
Nelson A. Rockefeller Empire State Plaza
Albany, New York 12234

April 1992

Paper presented at the 1992 annual meeting of the National Council on Measurement in Education

Copyright © 1992 by National Evaluation Systems, Inc. (NES®). All rights reserved.

"NES®," "National Evaluation Systems, Inc. ™," and the "NES®" logo are trademarks of National Evaluation Systems, Inc.

1. The authors would like to acknowledge Edward Murphy, Division Director of Assessment Product Services at National Evaluations Systems, for his work on the Professionalization of Teaching section of this paper.

THE NEW YORK STATE TEACHER CERTIFICATION EXAMINATIONS (NYSTCE)

New York State Assessment of Teaching Skills: A Multiple Method Assessment of Teacher Competency

Overview

A central part of the New York State Teacher Certification Examination Program is the Assessment of Teaching Skills. This facet of the examination program has two separate assessment components. The first component, written assessment, is required of all teacher candidates for provisional certification prior to entry into the profession. The second component, performance assessment, is required of all New York State teachers seeking permanent certification (required within five years of receiving a provisional certificate).

There is a growing recognition that the teacher stands at the center of education and that qualified teachers possessing a defined body of knowledge and skills play a central role in the educational process. The Assessment of Teaching Skills (ATS) for the New York State Teacher Certification Examination Program is an affirmative step in recognition of the need for qualified teachers that possess the necessary knowledge and skills to function effectively in New York State Public Schools.

The Professionalization of Teaching

Changing Views of the Teacher. Educational reform has focused attention on virtually every aspect of the complex process by which students acquire learning--school personnel, school environments, the students themselves, their parents. In recent years, significant attention has been brought to bear on the teacher; the teacher is emerging as a central figure responsible for the course of a child's education. While teachers have always been considered key players in the classroom drama, conceptions of exactly what role the teacher plays have changed dramatically over the years, moving from moral exemplar in the early years of the century through intellectual disciplinarian, nurturing caretaker, and technological delivery system in later decades. The latest conception of the teacher's role can best be conceptualized as the professionalization of teaching.

The Professional Teacher. Over the past decade, a view of teachers as respectable members of a challenging and worthy profession has been emerging and gathering strength both within the teaching community and outside it. This view presents teaching as a highly skilled activity with its own distinct body of knowledge. This body of knowledge is viewed as having to be diligently acquired by means of rigorous preparation through both college preparation courses and on-the-job experiences guided by structured interactions with mentoring co-professionals. This new view also presents classroom teaching as a complex, sophisticated set of skilled practices that requires an exceptional person to perform effectively.

While this view of teaching is not entirely new, it is clear that one of the most influential catalysts to the emergence of this viewpoint was the 1986 report of the Task Force on Teaching as a Profession, A Nation Prepared, which was itself a reaction to A Nation at Risk, the powerful and critical report of the National Commission on Excellence in Education, published in 1983.

The conception of teaching in A Nation Prepared encompassed several important features, among which were the following:

- the view that teachers must be highly qualified individuals rigorously prepared in their academic and professional disciplines
- the view that teachers must accept high standards for themselves and real accountability for their students' performance
- the view that teachers must be accorded the respect and freedom to design their own ways of achieving student success and other professional goals
- the view that teachers must be given a strong voice in school decision making and governance
- the view that salary and career incentives must be adequate to support the emergence of teaching as an attractive profession
- the view that teachers must be afforded a working environment and working conditions that are conducive to professional practices

The National Board for Professional Teaching Standards. A specific recommendation of A Nation Prepared that may prove to be quite significant for the teaching profession was its call for the creation of a National Board for Professional Teaching Standards "to establish high standards for what teachers need to know and be able to do, and to certify teachers who meet that standard." In 1987, that Board was created and began to undertake the task of establishing and maintaining high standards for a newly emergent teaching profession. The work of the NASCENT Board, together with a research base that focused on effective teaching and teacher expertise, has strongly affected the current conceptualization of teaching and the teacher.

In its published report, Toward High and Rigorous Standards for the Teaching Profession (NBPTS, 1989), the Board describes five "core propositions" that will guide its definition of a professional teacher. These propositions are presented here.

1. Teachers are committed to students and their learning.
2. Teachers know the subjects they teach and how to teach those subjects to students.
3. Teachers are responsible for managing and monitoring student learning.
4. Teachers think systematically about their practice and learn from experience.
5. Teachers are members of learning communities.

The portrait of the skilled teacher that develops from these propositions is the polar opposite of the conception that has dominated the research literature of the '60s and '70s and much of recent criticism of teachers. The teacher emerges as a knowledgeable, thoughtful professional who is skilled in the intricacies of subject matter, pedagogical methods, classroom management, and human relations. The teacher becomes the firm but flexible center of a dynamic swirl of purposeful activity aimed at helping

students learn. Teachers are seen as constant learners working together to perfect their practice. Above all, they are regarded as proud, competent individuals firmly in control of their classrooms and all the complex interactions that occur in them.

The Teacher Assessment Project. In the mid-1980s, the Teacher Assessment Project, a group of researchers formed around Stanford University professor Lee Shulman to investigate issues in the assessment of teachers and teacher candidates, became involved in research work that accorded with a professional view of teaching. Shulman and Sykes (1986) underscored the central theme of the professionalization viewpoint:

> Teaching has a substantial and growing knowledge base that incorporates but is distinct from the general norms of civility, from the general ethics of child-rearing, from the disciplines that are dealt with as subject matters in the school curriculum, and from the social sciences employed in educational research and practice. A large core of that knowledge base is amenable to the formation of persuasive and testable standards that, in part, define a profession.

The Teacher Assessment Project (TAP) thus supported the growing viewpoint that teaching was a skilled, knowledgeable act. In its major contribution to the definition of the knowledge base of teachers, the TAP attempted a formulation of the several kinds of knowledge potentially possessed by teachers. A 1986 list, which Shulman and Sykes presented as a "starting point for deliberation," offers the following summary of these kinds of knowledge:

1. General/Liberal Education, including basic skills of reading, math, writing, and reasoning.
2. Content knowledge in the domains in which teaching will occur.
3. Content-Specific Pedagogical Knowledge
4. General Knowledge of Pedagogical Principles and Practice
5. Curricular Knowledge
6. Understanding of Student Diversity and Individual Differences
7. Performance Skills (including voice, manner, poise)
8. Foundations of Professional Understanding (including history and policy; philosophy and psychology; cultural and cross-cultural factors; professional ethics)

In this formulation, Shulman and his group clearly extended the teacher effectiveness notion of generally useful and applicable principles of pedagogical practice to encompass the idea that a somewhat generic pedagogical reasoning process might be applied to each specific subject-matter area, forming a unique body of practice that represented the intersection of pedagogy and content. This notion was further described and particularized in Wilson, Shulman, & Richert (1987), focusing particularly on the fruitful notion of content transformation and reflective practice. The research tradition established by the Teacher Assessment Project has captured the research agenda at many institutions, including Michigan State University and the University of Pittsburgh.

Research on Expertise. Another strand of research that has informed the reconceptualization of teaching is the study of differences in thought and performance manifested by novices and experts. The notion underlying this body of research is that the habitual practices of experts in many realms of behavior, which may appear to the uninitiated to be "good instincts" or "natural talent," are largely the result of a knowledge base that is superior in terms of both content and organization (Berliner, 1989). What appear to be rational, unreflective, and nondeliberative responses in challenging situations are in fact the product of years of practice and subtle learning and reflection. While Berliner casts some doubt on the ability of expert teachers always to verbalize their underlying knowledge bases, the work on expertise has greatly influenced the new teacher assessment movement.

The differences in performance between novice and experienced teachers are reflected in the design for the New York State ATS. Two separate assessments are provided, one for the novice teacher (provisional-written assessment) and one for the more experienced professional (permanent-performance assessment).

THE NEW YORK STATE ASSESSMENT OF TEACHING SKILLS APPROACH

The New York State Assessment of Teaching Skills recognizes teachers as professionals and assesses teachers accordingly. The New York State Assessment of Teaching Skills stands in the line of research and conceptualization leading toward teacher professionalization that has been traced in this paper. The principles of professional practice in instruction and classroom management, reflective consideration of effectiveness, ongoing learning in both formal and informal settings, and qualitative performance growth are central to professional teaching. In addition, extra-classroom responsibility (including-- working with other professionals, working with parents and the community, contributing to curriculum development and instructional improvement, and working with school administrators) is included in the ATS definition of the professional teacher role.

The professional teaching literature provided a solid foundation for the ATS, but the ultimate definition of the ATS reflects the job requirements of the New York State teacher as defined by professional New York State teachers.

The New York State ATS Objectives. The ATS is based upon a substantive definition of the professional teacher, which has been accomplished through the careful crafting of a set of research- and practice-based teacher objective statements. The statements were derived in part from recent research on teaching and in part from the work of practicing New York State classroom teachers. The objective statements will be validated as appropriate for the New York State program with substantial input from New York State educators. These statements, which will be made public to all examinees and the teaching profession through broad dissemination efforts, are designed to form the basis of all assessments for both components of the ATS.

The image of the teacher that emerges from the ATS objectives is of the same active, astute, concerned, flexible, reflective, and informed practitioner that has been emerging from the professionalism literature.

Examinees' knowledge and skills in four domains will be assessed by the ATS. While teaching is a unified activity, the components of teaching skills are divided into domains for assessment purposes.

I. Knowledge of the Learner. Understanding human development factors relating to the home school and community that may affect the learner, student diversity and learning processes.

II. Instructional Planning and Assessment. Understanding curriculum development instructional planning processes and formal and informal assessment.

III. Instructional Delivery. Understanding multiple approaches to instruction, motivational principles and practices, communication, and classroom management.

IV. Professional Environment. Understanding reflective practice, school-home relationships, teacher/student rights and responsibilities, and the structure of New York State schools.

GOALS OF THE ATS

The ATS is designed to meet several goals. The assessment is designed to meet traditional measurement requirements and at the same time is designed as an assessment that will be credible to the teaching profession. The assessment should be fair to candidates and be feasible to develop and administer.

Professionally Credible. The ATS assessment should have the support of the teaching profession. The ATS will measure authentic knowledge and skills defined by New York State teachers. The assessments of those skills will in turn be substantially defined by members of the New York State teaching profession. By defining and measuring real world educational needs as defined by New York State educators, the ATS can effectively achieve the support of New York State teachers.

Technically Sound. The ATS assessment should meet traditional measurement requirements. The assessment should be sufficiently valid and reliable to achieve the goals of the assessment. The ATS will be designed to effectively measure the domain of important knowledge and skills required of New York State teachers to perform effectively in New York State classrooms. Moreover, the assessment will be designed to reliably evaluate those requirements to allow a reasonable degree of confidence in the stability and consistency of the results.

Fair. The assessment should be fair to candidates. The ATS is designed to minimize the effects of extraneous factors on the results obtained; to the extent possible only those aspects of teaching performance targeted by the ATS should affect the decision made. Appropriate measures will be taken throughout the development process to maximize fairness.

Feasible. While many assessment approaches are available, the ATS should be designed to achieve the desired assessment goals in a way that is feasible. The assessment should be able to accommodate large numbers of examinees seeking certification under high stakes conditions. Moreover, the assessment should be developed and administered at a cost that is reasonable.

DESIGN OF THE ASSESSMENT OF TEACHING SKILLS

The assessment of teaching skills will employ a two-tiered approach: a written assessment at the entry-level and a performance-based assessment later in the teaching career. The written assessment, required for provisional certification, is designed to measure the knowledge and skills expected of teachers at entry into the profession. The performance-based assessment will be required for permanent certification (required within five years of receiving a provisional certificate). This assessment will be targeted at the performance level expected of a more seasoned teacher with several years of teaching experience and will be designed to tap this level of performance within the specific context in which the teacher is practicing. Separate assessments of teaching skills will be developed for elementary and secondary teachers to allow for variation in the teaching requirements at the two levels.

Written Component

Overview. The candidate for a provisional New York State certificate will be required to pass the written assessment. The written component of the New York State Assessment of Teaching Skills is designed to evaluate the candidate's knowledge of teaching skills. It is anticipated that the written component of the assessment of teaching skills will be targeted at all four domains of knowledge defined earlier.

Assessment Methods. The written component will employ several assessment approaches and both selection-type and supply-type items. A variety of machine-scorable item formats requiring the examinee to select from among four response options in response to a question will be included (e.g., clustered stimuli, stand-alone, K-type). Examinees will also be asked to write a response to a written assignment calling for the examinee to analyze an educational issue, provide a work sample in response to a given assignment or provide an appropriate solution to an educational problem presented.

Test Item Formats. The Assessment of Teaching Skills is designed to assess a broad sample of examinee knowledge related to the set of objectives that have been validated as important for an educator in New York State. Test questions on the written test may require knowledge regarding:

- a fact,
- a concept,
- a principle, or
- a teaching skill;

or may require the ability to apply or analyze:

- factual knowledge,
- a concept,
- a principle, or
- a skill in a given context.

The item formats that may be used on the written test are briefly described below.

Single Problem Items. A single question is asked. Examinees select the one best answer from a set of responses. The question may be presented directly or through reference to a situation, one or more excerpts, or one or more graphic stimuli.

Clustered Problem Items. Two or more questions are asked about a situation, an excerpt, a set of excerpts, or one or more graphic stimuli. Examinees select one answer to each question from a set of responses.

Multiple-Correct-Answer Items. A single question is presented directly or through reference to a stimulus. A list of several possible answers to the question is presented, and each possible answer is preceded by a roman numeral. More than one of the possible answers may be correct. Examinees select from four options the *set* of answers that is best (e.g., I and III only; II, III, and V only).

Written Assignment. A prompt or assignment asking examinees to analyze an educational problem, produce a sample of their work, or solve an educational problem is presented and examinees respond in writing.

Test Length. The written component will contain approximately 80-100 machine-scorable items and a single written assignment.

Test Administration. The written component of the New York State Assessment of Teaching Skills will be group administered in a single morning or afternoon session of approximately 4 hours in length, although it is anticipated that the typical examinee will complete the examination within 3 hours; the assessment will be designed such that time will not be a significant factor in examinee performance. Examinees will be permitted to work at their own pace completing individual questions and test sections in the order they deem appropriate; individual sections will not be timed.

Examinees will be provided with a test booklet containing the machine-scorable questions and the written assignment along with appropriate instructions for completing the examination. An answer document with sections for completing both the machine-scorable and written assignment sections of the test will be provided.

Scoring. The machine-scorable portion of the examination will be recorded by examinees on an answer sheet that will be optically scanned following the administration. The examinees' written response will be evaluated independently by two New York State educators on a predetermined scale using a focused holistic scoring methodology. The focused holistic approach requires the scorer to make a single, overall judgment about the quality of the examinee's response based on an analysis of identified dimensions of interest derived from the objectives defining the examination and the scoring criteria established. Examinees will be scored on the basis of teaching skills, not subject area knowledge. Candidates for an elementary certificate will be scored by elementary educators and candidates for a secondary certificate will be scored by secondary educators. For any response where two scorers disagree substantially on the score given to a written response, additional scoring will be required to evaluate the response to help to determine a final score.

Performance-Based Component

The performance-based component of the New York State Assessment of Teaching Skills is designed to assess prospective teachers' ability to apply teaching skills. The New York State Department of Education is currently exploring the feasibility of using a video-based performance assessment for this purpose. Tentative plans for this component are described below.

Examinees will be asked to demonstrate their teaching skills which they will videotape for scoring purposes. It is anticipated that the teaching skills video-based component will be targeted primarily at Domain III (Instructional Delivery), although other domains may be assessed as appropriate. Teachers seeking permanent certification will be evaluated based on their ability to apply multiple approaches to instruction, apply motivational principles and practices, communicate effectively, and effectively manage the classroom as called for in Domain III.

Assessment Methods. A video-based observational approach may be employed to assess New York State educators' teaching skills. The teacher would be asked to teach a lesson of the teacher's choosing to his or her own students. The teaching episode would be videotaped by the teacher according to specifications to be provided. The candidate would provide about two pages of written information about the teaching episode as prescribed by the assessment process. This written information would include information about the students, the objectives of the lesson, and a reflective statement.

Test Length. A single videotaped episode of instruction would be required of the examinee. Teachers will be instructed to limit the episode to approximately 30 minutes.

Test Administration. Examinees would be provided a preparation manual describing the assessment, necessary forms to be completed along with the assessment, the specifications for the assessment, and instructions for completing the assessment. Examinees would be responsible for obtaining the videotape equipment, providing a blank videotape and videotaping their episode. The examinee may choose the day, time, and content for the videotaped episode within the constraints of the assessment.

Scoring. The videotaped episode would be submitted for scoring along with the verification of identification. Video responses would be scored at one of the scheduled scoring sessions to be held at a New York State site where certified New York State educators would score the responses. The examinee's video response would be evaluated independently by two New York State educators on a predetermined scale using a focused holistic scoring methodology. The focused holistic approach requires the scorer to make a single, overall judgment about the quality of the examinee's response based on an analysis of identified dimensions of interest derived from the objectives defining the examination and the scoring criteria established. Examinees would be scored on the basis of teaching skills, not subject area knowledge. Candidates for a permanent elementary certificate will be scored by elementary educators and candidates for a permanent secondary certificate will be scored by secondary educators. For any response where two scorers disagree substantially on a score given to a videotaped episode, additional scoring would be required to determine its final score.

VIDEO-BASED PERFORMANCE STATUS

Video Feasibility Study. Over the past several months, National Evaluation Systems (NES®) and the NYSED have been conducting a study to determine the feasibility of using a video-based performance assessment. The first component of the study has focused on the technical quality and equipment issues associated with video production. The second component of the study has focused on the procedures for having teachers tape their own classroom performance given a set of instructions and the procedures for scoring the teacher-produced video tapes.

While the decision of whether or not to proceed with the video-based performance assessment has not yet been made, it appears that this approach has promise. Teachers, in general, are able to produce a videotape of their classroom performance that appears scorable in a standardized fashion. Audio and video quality appear acceptable using typical video equipment found in public schools. Nevertheless, there are still several issues that remain to be addressed. There are both technological and implementation issues that will be considered in the coming months.

Technology Issues. Some of the technological issues being explored include:

- camera placement
- camera movement
- audio/video quality
- equipment access
- optimal v. typical performance
- scorer training and scorer qualification

Implementation Issues. Some of the implementation issues being explored include:

- teacher confidentiality
- parental permission to videotape students
- school/administrator cooperation
- impact of removing failing candidates from the classroom
- impact on teacher education
- verifying candidate identify

REFERENCES

American Educational Research Association, American Psychological Association, & National Council on Measurement in Education. (1985). *Standards for educational and psychological testing*. Washington, DC: Author.

Berk, R. A. (1982). Introduction, In R. A. Berk (Ed.), *Handbook of methods for detecting test bias*. Baltimore, MD: Johns Hopkins University Press.

Berk, R. A. (1980, April). *A consumer's guide to criterion-referenced test "reliability."* Paper presented at the annual meeting of the NCME, Boston, MA

Berliner, D. C. (1989) The development of expertise in pedagogy.

Berliner, D. C. (1986). In pursuit of the expert pedagogue. *Educational Researcher, 15*, (pp. 5-13).

Cronbach, L. J. (1971). Test validation. In R. L. Thorndike (Ed.), *Educational Measurement* (pp. 443-507). Washington, DC: American Council on Education.

Equal Employment Opportunity Commission, Civil Service Commission, U.S. Department of Labor, & U.S. Department of Justice. (1978). Adoption by four agencies of uniform guidelines on employee selection procedures. *Federal Register, 43*, 38290-38315.

Holmes Group. (1986). Tomorrow's Teachers. East Lansing, MI. Author.

Huynh, H. (1976). On the reliability of decisions in domain referenced testing. *Journal of Educational Measurement, 13*, 256-64.

Livingston, S. A. (1972). Criterion-referenced applications of classical test theory. *Journal of Educational Measurement, 9*, 13-29.

Messick, S. H. (1989). Validity. In R. Linn (Ed.), *Educational Measurement* (3rd ed.) (pp. 13-92). New York: Macmillan Publishing Co.

National Board for Professional Teaching Standards. (1989). *Toward high and rigorous standards for the teaching profession: initial policies and perspectives of the National Board for Professional Teaching Standards*. Detroit: NBPTS.

National Evaluation Systems, Inc. (1991). *Bias issues in test development*. Amherst, MA. Author

Nunnally, J. C. (1978). *Psychometric theory*. New York: McGraw-Hill.

Plake, B. S. (1980). A comparison of a statistical and subjective procedure to ascertain item validity: One step in the test validation process. *Educational and Psychological Measurement, 40*, 397-404

Sandoval, J., & Miille, M. P. W. (1980). Accuracy of judgments of WISC-R item difficulty for minority groups. *Journal of Consulting and Clinical Psychology, 48,* 249-253.

Shepard, L. A. (1982). Definitions of bias. In R. A. Berk (Ed.), *Handbook of methods for detecting test bias.* Baltimore, MD: Johns Hopkins University Press.

Shulman, L. S., Sykes, G. (1986). A national board for teaching: In search of a bold standard. Paper prepared for the *Task Force on Teaching as a Profession, Carnegie Forum on Education and the Economy*

Task Force on Teaching as a Profession. (1986, May). *A nation prepared: teachers for the 21st century.* Washington, DC: Carnegie Forum on Education and the Economy.

Tittle, C. K. (1982). Use of judgmental methods in item bias studies. In R. A. Berk (Ed.), *Handbook of methods for detecting test bias.* Baltimore, MD: Johns Hopkins University Press.

University of the State of New York, The State Education Department Office of Teaching. Amendments to the Rules and Regulations of the State of New York, Regulations of the Commissioner of Education. Pursuant to *Sections 207, 305, and 3004 of Education Law.*

Wilson S. M., Shulman L. S., Richert, A. E. (1987). One hundred fifty different ways of knowing: Representation of knowledge in teaching (edited by J. Calderhead). *Exploring teachers' thinking.* Sussex, England: Holt, Rinehart and Winston.